Conformational Searching Methods for Small Molecules. III. Study of Stochastic Methods Available in SYBYL and MACROMODEL

A. M. TREASURYWALA,* E. P. JAEGER, and M. L. PETERSON
Sanofi-Winthrop Inc., Collegeville, Pennsylvania

Received 7 March 1995; accepted 19 September 1995

ABSTRACT

In a continuing effort to provide the computational community with a reference work comparing all of the available conformer searching methods, we have exposed the standard set of small molecules to two commonly used stochastic searching techniques. Advantages and limitations of these methods are discussed. © 1996 by John Wiley & Sons, Inc.

Introduction

The importance of finding low energy conformers for the success of any molecular modeling project cannot be overstated.[1] This fact, coupled with the plethora of available methods to search conformational space and the paucity of reference works that compare them,[2-4] prompted us to undertake a survey of the commonly used methods. Previous comparisons have not exposed the same molecules to a comprehensive collection of searching techniques or have done so with only a single molecule or a small number of them. Our intent was to expose a set of molecules that represented a variety of common functionality and scaffolding found in many organic problems to as wide a selection of searching methods as possible. The object has been to expose the fixed set of molecules to each of the methods as they are most likely to be used by the end user (i.e., placing a heavy emphasis on the default conditions prescribed by the authors), and then to tabulate and analyze the findings so as to provide a database of results that will enable users to choose the method that seems most practical to them for their particular problem.

It should be emphasized at the start that it is not the intent of this survey to find the method that will work all of the time or even most of the time. Nor is it the intent in this first pass to find the optimal conditions for running a given method for a given set of molecules. These are all worthy objectives and may be pursued at a later stage by

*Author to whom all correspondence should be addressed at Allelix Biopharmaceuticals, Inc., 6850 Goreway Drive, Mississauga, Ontario L4V 1V7, Canada.

CCC 0192-8651 / 96 / 091171-12

us or by others as the need arises. Problems differ widely in the amount of "completeness" that is needed and the type of answers that are sought. For example, most of the conformational analyses that we have carried out in practical drug design problems are more concerned with quick and reliable ways of deciding what "low energy conformer" means for a given molecule than in finding the global minimum. This is because the conformer that fits a particular enzyme pocket or

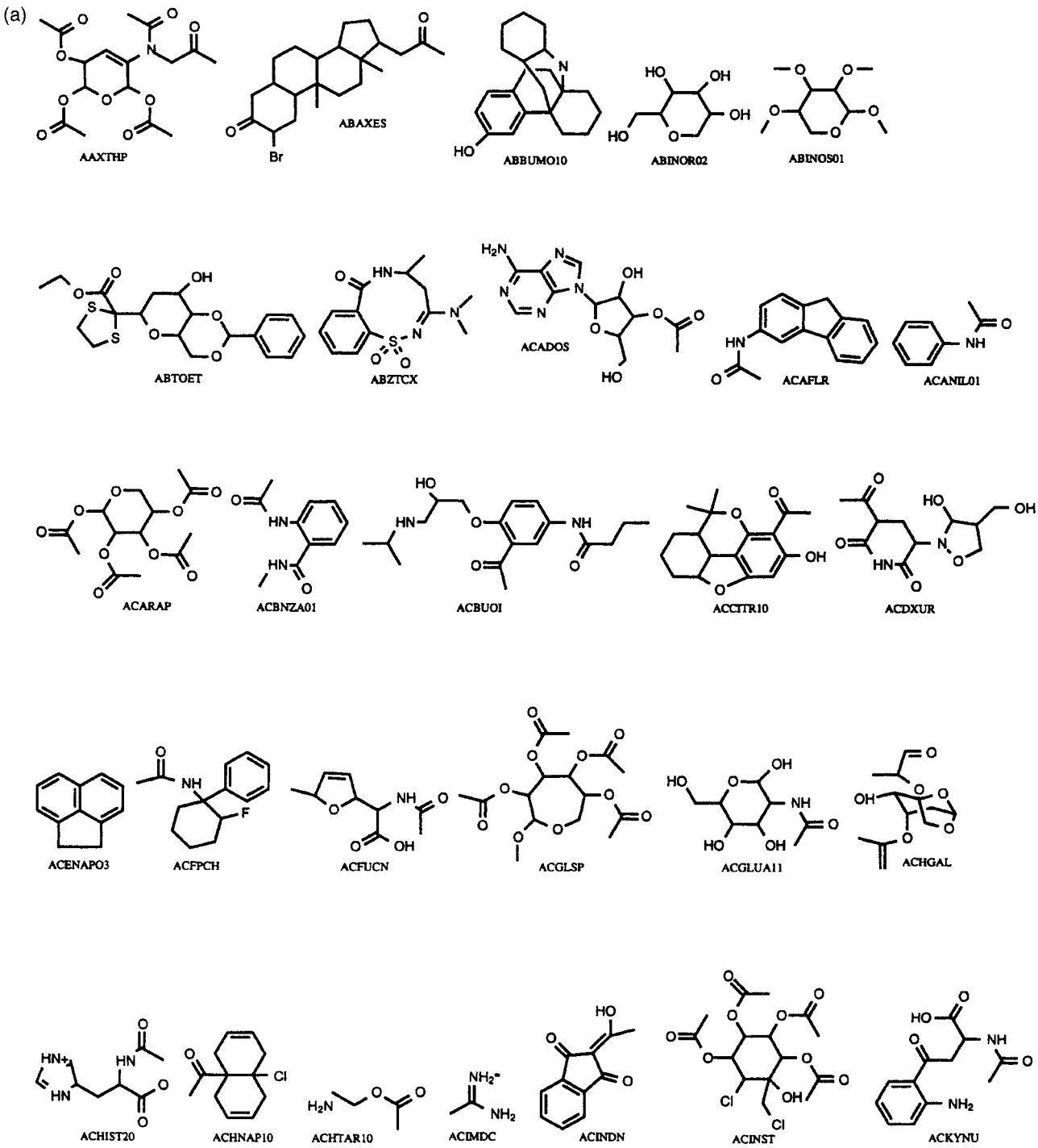

FIGURE 1. Structures of the molecules used in this study.

(b)

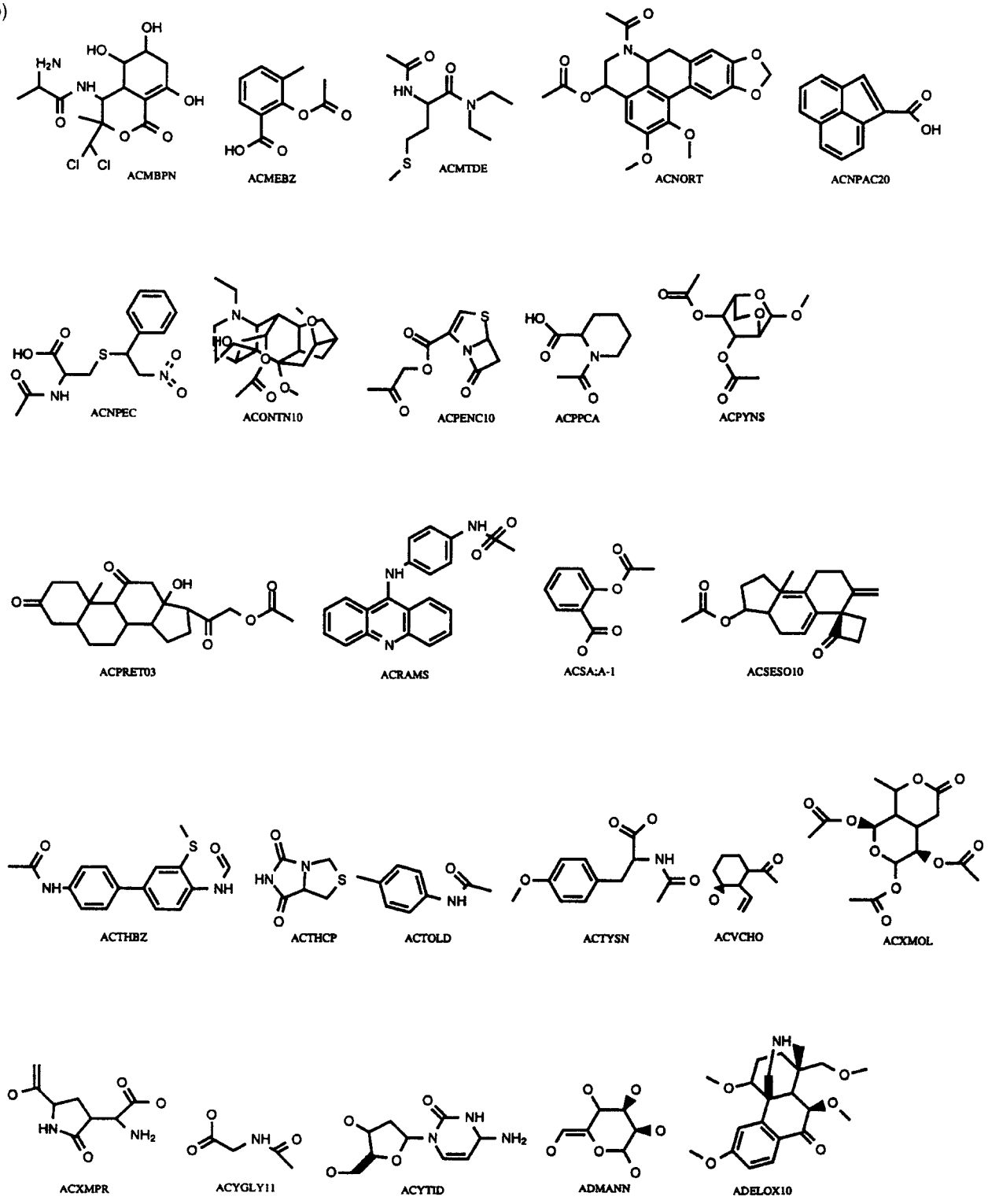

FIGURE 1b. (continued)

pharmacophore requirement can easily be determined. After it has been found, however, the question arises about how much energy needs to be invested in the molecule to get it from the conformers that exist (say in solution) to this "biologically relevant" conformer. If this number is less than a cutoff (usually arbitrarily taken to be the strength of a hydrogen bond by us or roughly 3 kcal/mol) then it is assumed that this biologically relevant conformer can be reasonably at-

(c)

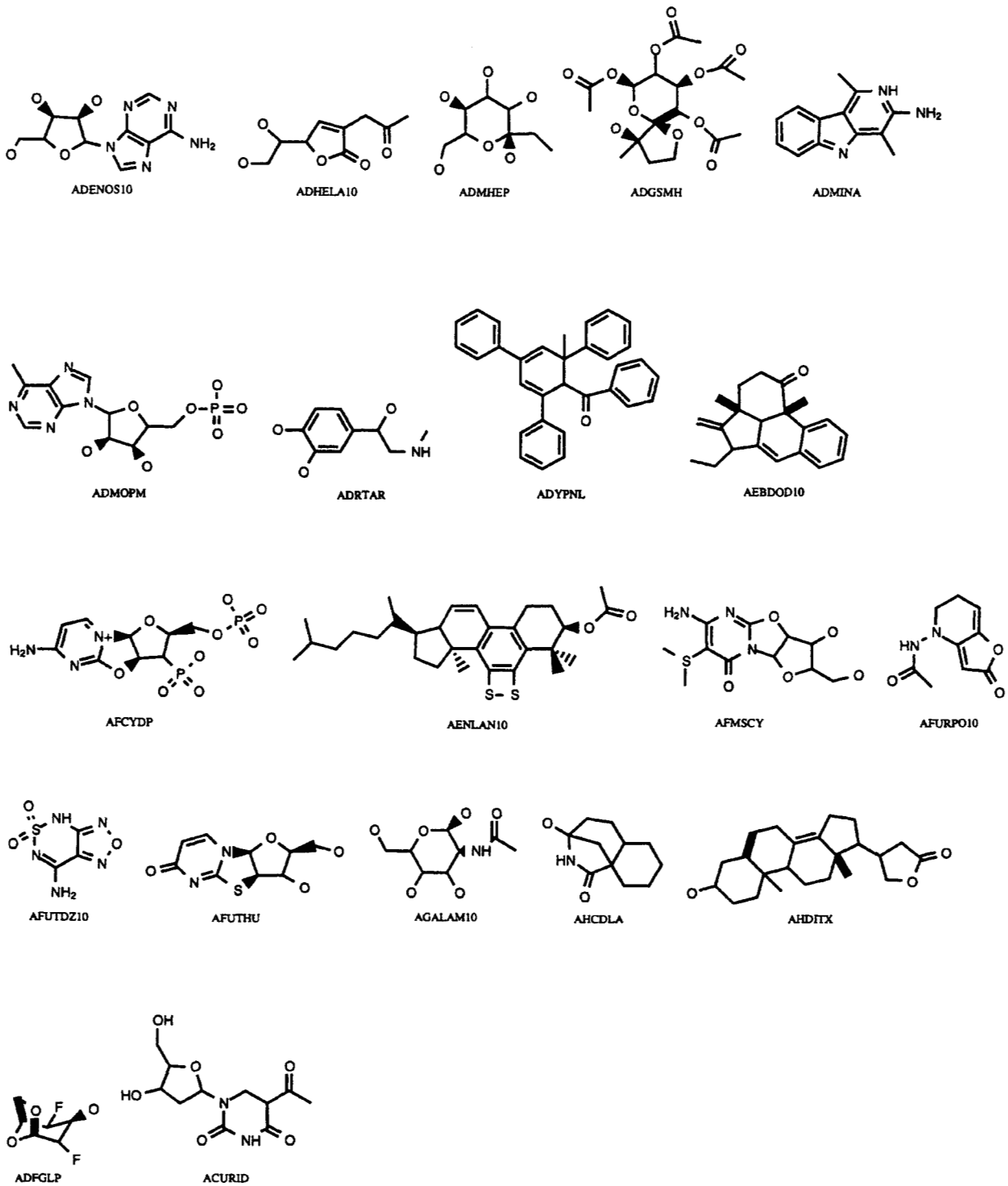

FIGURE 1c. (continued)

tained and that the molecule therefore "supports" the developing hypothesis. If, however, the conformer required to satisfy the hypothesis at hand is "very" energetically unfavorable, then it may be appropriate to question the hypothesis that requires it. There are, however, other types of problems for which the quest for the global minimum is of the utmost importance. Yet other instances require the enumeration of ALL of the possible low energy conformers. One example of the latter requirement would be in the field of NMR interpretations. Naturally, investigators in any of these situations will want to know the strengths and weaknesses of the different available methods so that they may best decide for themselves which method best suits their current needs. While the data presented in these reports may be of use to those seeking methods that assure one of finding the global minimum, they will not *per se* explicitly provide this information. In many of the practical cases selected for this survey the so-called global minimum is indeed unknown.

Our approach has been to subject a preselected set of molecules to the method being examined and to ask some standardized questions of the results. (1) How did the energy of the "lowest energy conformer" found compare (after it was minimized to convergence) to that of the "reference conformer"? The reference conformer is defined here as the one that resulted when the same force field was used to minimize the crystal structure to the same level of convergence. (2) Was the reference conformer actually visited during the search? The closest conformer to the reference one was identified by two separate criteria: (a) the overall heavy atom RMS value of this conformer when compared to the reference conformer and (b) the torsional RMS (see definition in previous studies[5,6] and references cited therein[7]) value of this conformer when compared to the reference. In many cases, the minimum of these two values was not found in the same conformer. In such cases each conformer was examined further. Thus, any search was simplified by examining the following four conformers: the reference one, the lowest energy one, the heavy atom RMS one, and (if different from the previous) the torsional RMS one. We recognize that this method of analysis of the result is somewhat arbitrary but it was considered adequate given the questions that we had set out to ask in this survey.

The first article in this series examined the SYBYL search method as an example of the systematic searching technique which is commonly available in several commercial packages.[5] In this study we set up the general paradigm and the criteria that would be used for the evaluation of each of the methods. In the second study, we examined the performance of a method that was developed in-house and that relies on the use of the genetic algorithm for optimizing the energy of a population of conformers of a given molecule.[6] Both methods used the same molecules. In this article we apply two examples of the category of so-called stochastic methods, the random method in SYBYL[8] and the Monte Carlo method in MACROMODEL,[9] to the same set of molecules. The results are deliberately analyzed in as similar a way as is possible to the methods used in previous studies in this series so as to make the comparisons as direct as possible.

The stochastic methods used in conformational analysis differ from the systematic methods in several basic ways. At the heart of the former class of methods is the strategy that a molecule will be randomly perturbed from a "suitably chosen" starting conformer and then minimized. The new conformer thus formed will be evaluated for two properties: (1) Was it observed before? and (2) What is the energy of the minimized structure? It has been claimed[7] that the number of times that a conformer is discovered is related to the population of the particular minimum at equilibrium. It has also been estimated that the chances of finding new minima are negligible after the found ones have all been visited at least six times. Thus, most random methods will attempt to perform a random perturbation until all of the found minima have been visited some predetermined number of times or until they run out of some arbitrarily large number of trials. This latter parameter is usually set based upon the amount of computer (and also wall-clock) time available to the investigator. We would usually consider a run to have stopped prematurely if the reason for its stopping is because of the latter termination criterion. In the strictest sense, no conclusions could be drawn about the relative populations of the various minima under such circumstances.

The choice of which conformer to apply the given perturbation to is a major differentiating point for the various methods as is the type and extent of perturbation performed. Some sample all of the found conformers in a way that is proportional to their energy while others sample them based on which ones were used *least* in previous attempts. This latter method favors the exploration of new parts of the conformational landscape.[10]

Finally, some methods perturb all of the cartesian coordinates of the molecule and then minimize the resulting strained conformation. The strain acts as a form of "reservoir" of potential energy during the minimization and, depending on the force field and minimization technique used, may result in novel conformers being sampled. Even those methods that perturb torsional angles do so in different ways. The two methods being examined in this study deal with the torsional list differently. The SYBYL-RANDOM method perturbs all of the torsions each time. However, the MACROMODEL-MC implementation perturbs only a subset of the list that is randomly chosen. Other implementations have been reported as well.[10] The stochastic methods are generally regarded as faster but less thorough than the so-called systematic methods.[1] However, examples of the thorough examinations of this generally held axiom are scarce. It was therefore of interest to us to determine the extent of the validity of these two generalizations based on real data. Systematic search methods always require a choice of the number of parameters to be varied and the steps to be used. If there are too few parameters, or if the steps are too large, many conformers will be missed. If the number of parameters is increased and the size of the steps is decreased, the number of conformations evaluated, and hence the time required, explodes. A limiting feature of the systematic methods is that the optimal choices of parameters and step sizes vary from case to case and no method is available for predicting them in advance.

Methods

As has already been pointed out, the molecules are the same standard set that has been used throughout this series of studies. They have been chosen for two principle reasons: (1) they represent the typical functionalities that are commonly found in small molecules of interest to us and so are of some practical value; and (2) they have been used before in other independent studies which have validated the various force fields that are used in these studies.[11-14] Thus, any discussions about the force fields themselves and the role that their actual validity may play for the results reported here are considered outside the scope of this article.

The molecules were all extracted from the Cambridge Structural Database (CSD) and were prepared for this study as follows. The CSD entry was reduced to a single molecule in each case by removing the solvent molecules or other molecules that may have been included in the original crystal structure. These single-molecule entries were atom typed in SYBYL6.03. Consistency in the atom types was considered to be of greater importance than the actual choices made for a given atom and was maintained both within a given study from one molecule to another as well as between studies in this series to the extent possible. Naturally, when a translation to acceptable atom types for another program was called for, the same care was used in making these translations. They were chosen to be consistent with the bonding environment of the atom and are, to the extent possible, consistent with those used in the previous studies.[5,6] The bonding used was likewise chosen to be consistent with the atom types and with previous work as much as possible. After the runs were completed, all of the analyses were performed in SYBYL. This step ensured that all runs were analyzed in exactly the same way after they were completed. In cases in which backtranslation to SYBYL atom types was needed before an analysis could proceed, the original mol2 file was used and only the coordinates of the atoms of the final molecular conformer as supplied by the program being used were substituted. This ensured that the numbering and typing of the atoms was completely consistent during the analysis.

All exocyclic bonds except those at the ends of aliphatic chains (i.e., leading to terminal methyl groups) were chosen as rotatable. All other bonds were not modified from the torsions that were present in the original crystal structures. Amide bonds in this study were also held fixed. Ring conformations have been held fixed throughout this study as they were in the two previous studies. The reasons for this choice were: (1) it was felt that the addition of ring conformations to the overall search would serve simply to add rotatable bond numbers to the overall searching method and would not necessarily add any new insights; and (2) ring searching methods could be examined in a separate study comparing all of the available methods. This has been done in the case of cyclododecane by others.[4]

Each of these molecules was minimized to convergence (gradient) in the force field that was to be used in the study. This final resulting structure was the starting point for the conformational searching. It is important to note this step here. None of the force fields used so far recognized the original crystal structure as a minimum. Thus,

during minimization to convergence, these force fields typically modified the coordinates in different ways. For example the force field in SYBYL[11] tended to modify the bond lengths and bond angles more than it did the torsion angles to relieve the strain it perceived in the crystal structure. However, MACROMODEL tended to modify the torsion angles as well, and therefore produced conformers that were sometimes quite different from the starting crystal conformer during this step. A discussion of the relative merits of these two force fields is outside of the scope of this article. However, we felt that a proper analysis of the performance of the conformational searching method would be obtained only if we compared the results obtained using a given force field with the conformer in which that same force field would place the starting crystal structure. This was the main reason for performing the analysis in the way it was done. Understandably, it may appear to some that the results of a search were judged by comparison with different reference conformers in the two methods. It was our assessment that this was a more even comparison of the two searching methods in the long run.

SYBYL-RANDOM METHOD

All of the SYBYL-RANDOM runs were performed on a Silicon Graphics Onyx VTX machine with eight processors. The program was set to check for and avoid the inversion of chiral centers. The energy cutoff was set at 70.0 Kcal/mol. All conformers within this initial energy value from the lowest that was found at that particular instant in the search would be included. The RMS threshold was set at 0.200. Any conformer that was within this value in its heavy atom RMS to one that had already been found would be included in the list as another occurrence of that found conformer. Any acceptable conformer that was outside of this cutoff would be listed as a new conformer. The maximum number of search iterations was set to be 1000 for this study and the maximum number of hits was set to six. If the method found all of the identified low energy conformers six times *before* it had exhausted the maximum number of trials permitted (1000) then it would terminate normally. If the method had made 1000 attempts to find "low energy" conformers, and had not found all the ones it had already identified six times or more, then it would also terminate the run. The gradient convergence was set to 0.005 for termination of the minimization operation and the maximum number of iterations during a minimization was set to 100. The results of runs on each of the molecules are given in Table I.

MONTE CARLO METHOD IN MACROMODEL

All of the MACROMODEL Monte Carlo conformational search runs were performed on a Vax 600 running Version 4.7 of the VMS operating system using Version 2.5 of MACROMODEL. The MM2* force field was selected for the energy computations. The program was set to run 100 Monte Carlo (MC) steps for each compound. Each MC step began with the starting geometry of the previous MC step and a maximum of two torsions were varied in each MC step. The trial conformation from each MC step was then subjected to minimization using a block diagonal Newton–Raphson minimizer. The minimization continued until a derivative convergence criterion of 0.1 kJ/Å · mol was reached or until a maximum of 250 iterations had been performed. Duplicate conformers were detected using a heavy atom RMS criterion of 0.25 Å. A trial conformer with a heavy atom RMS < 0.25 Å to an existing conformer was counted as another occurrence of that existing conformer. A trial conformer was counted as a new conformer if its heavy atom RMS was > 0.25 Å when compared to each of the previously existing conformers. The results of these MC runs are given in Table II.

Results and Discussion

Several interesting findings emerge from this study. First, there appears to be no simple correlation between the flexibility of the molecule, as measured by the number of rotatable bonds present, and the number of conformers that are found before termination. If, for example, one examines the results obtained for all molecules with five rotatable bonds, one finds that there are 11 such molecules in this study. They are ACDXUR, ACFUCN, ACHGAL, ACNORT, ACPRET03, ACPYNS, ACURID, ACXMOL, ADELOX10, ADYPNL, and AFMSCY. The number of acceptable conformers ranges from 21 (AFMSCY) to 325 (ADELOX) in the SYBYL-RANDOM-based searching method with one particular case (ACHGAL) finding no energetically acceptable conformer in

the first 1000 attempts. Although the two methods are nominally very similar the details of the results obtained from them often differ. Continuing the analysis of the 11 cases of five rotatable bonds, we find that the MACROMODEL results range from 48 (ACXMOL) to 557 (ADYPNL) acceptable conformers found. Not only are these high and low points not the same molecules in the two methods, but ADYPNL, for example, which represents the high point for MACROMODEL, is one of the lowest in SYBYL. Other counterexamples were also found. MACROMODEL also failed to find acceptable low energy conformers for one molecule (AFMSCY) but it was not the same molecule as the one identified by SYBYL (ACHGAL). For the latter molecule, MACROMODEL easily identified 132 low energy conformers. This example suggests two factors: (1) the methods are stochastic and thus, by their very nature, cannot assure the user of a complete search—although the probability of finding good low energy conformers by either method is very high as seen from the data. It may be prudent to perform the search from several starting conformers and to compare the results or alternatively to perform the search with as many different stochastic methods as possible so as to avoid generalizations from incomplete data. Either of these suggestions will of course considerably increase the time for the search. (2) As mentioned previously there is no simple way to connect the number of rotatable bonds and the number of low energy conformers that may reasonably be expected to be found. Another observation that emerged from this analysis is that the preset maximum number of searches needs to be increased as the number of rotatable bonds in the molecule increases. For example, none of the conformational searches for molecules containing five or more rotatable bonds actually terminated in SYBYL-RANDOM because this method found all of the low energy conformers six times. All of them terminated because they exceeded the number of attempts allowed (1000). It is fair to say that even some of the molecules with less than five rotatable bonds met with the same fate, but beyond five rotatable bonds this is apparently a very highly probable event.

The so-called energy window is used to determine if a new conformer will be considered energetically acceptable and therefore added to the collection of low energy conformers. For the runs reported it was set in each case to the default values recommended in the manuals for the software. However, this parameter was found to be a strong factor in determining the number of low energy conformers that were reported. If the window was reduced, fewer conformers were found to be acceptable and many more runs were terminated because they had found all enumerated low energy conformers the required number of times (six). However, more of them also found no acceptable conformers within the number of set attempts. Thus the results that were obtained for the group of molecules varied greatly with modest changes to this energy window parameter. The quantitative details of this result are not reported here for purposes of brevity. Therefore, it is prudent to examine the dependence of the results on this parameter before drawing other conclusions from them.

As Tables I and II show, there is a reasonable correlation between the energies found for all of the conformers that were singled out for study. The energy of the minimized X-ray conformer compared, in general, very favorably with that of the lowest energy conformer found by the technique. In several cases, the methods actually found conformers with lower energy than the crystallographic reference. As has been pointed out in our previous studies, this is not in any way a challenge to the method. The consistent result of finding only conformers considerably higher in energy could be a strong indictment against the method; however, was not found to be the case with either of the methods examined here. The same type of general correlations were found with the energies of the conformer with the closest torsional RMS or heavy atom RMS to the reference conformer. One exception in both of the latter cases was ACHTAR when examined using the SYBYL-RANDOM method. This molecule's reference conformer had an energy of 9.63 kcal/mol and indeed, the method found a conformer of energy 9.11 kcal/mol. However, the closest conformers that it found in terms of the torsional RMS and the heavy atom RMS were of energy values 50.65 and 62.36 kcal/mol, respectively. This means that the search did not find the crystal conformer within its low energy band of structures. MACROMODEL-MC, on the other hand, found the reference energy of this molecule to be 47.19 kJ/mol and indeed also found a conformer (#3) which was lower in energy (36.78 kJ/mol), however, it succeeded in visiting the reference conformer also (conformer #27) as seen by

TABLE I.
Results from the SYBYL-RANDOM Runs.

RESULTS FROM SYBYL'S RANDOM

NAME	Xraymin		E Best			Tor Best			RMS Best				
	Conf	Energy	Conf	Energy	TRMS	RMS	Conf	Energy	TRMS	RMS	Conf	Energy	TRMS	RMS	CPUtime
aaxthp2	768	32.39	277	26.36	85.31	1.65	409	29.37	7.19	0.44	409	29.37	7.19	0.44	39585.92
abaxes	8	33.66	2	31.90	76.37	0.62	1	32.82	6.09	0.12	1	32.82	6.09	0.12	2060.82
abbumo10	2	15.07	1	13.94	8.70	0.12	1	13.94	8.70	0.12	1	13.94	8.70	0.12	70.41
abinor02	2	1.74	1	2.28	63.51	0.08	1	2.28	63.51	0.08	1	2.28	63.51	0.08	14.56
abinos01	2	1.77	1	1.45	59.30	0.02	1	1.45	59.30	0.02	1	1.45	59.30	0.02	14.93
abtoet	150	10.12	70	6.67	72.91	1.70	134	7.04	25.57	0.64	134	7.04	25.57	0.64	56183.18
abztcx															
acados	195	51.09	15	47.31	71.96	0.62	93	48.55	58.50	1.41	124	48.99	82.45	0.60	27509.65
acaflr	9	14.98	1	13.90	1.87	0.09	1	13.90	1.87	0.09	1	13.90	1.87	0.09	562.88
acanil01	5	5.93	1	5.02	6.03	0.07	1	5.02	6.03	0.07	1	5.02	6.03	0.07	177.20
acarap	504	10.69	86	8.84	49.84	1.00	26	9.03	3.69	0.24	26	9.03	3.69	0.24	30923.62
acbnza01	23	8.69	1	7.32	109.03	0.90	22	9.74	39.20	0.70	18	7.39	90.39	0.13	2023.10
acbuol	770	11.99	378	6.40	103.36	3.35	215	7.73	64.85	1.97	75	7.31	69.30	1.23	63884.81
accitr10	3	22.02	1	17.91	8.60	0.11	1	17.91	8.60	0.11	1	17.91	8.60	0.11	212.06
acdxur	170	18.31	123	14.12	108.15	1.09	100	14.23	10.66	0.69	57	14.82	52.88	0.68	17901.51
acfpch	8	6.05	1	4.22	16.90	0.30	1	4.22	16.90	0.30	1	4.22	16.90	0.30	1021.61
acfucn	112	12.15	44	9.51	68.15	1.10	42	9.52	23.49	0.42	42	9.52	23.49	0.42	9276.12
acglsp	740	21.21	399	16.02	66.40	2.96	31	19.19	28.78	0.61	31	19.19	28.78	0.61	44683.95
acglua11	39	4.18	1	2.51	79.82	0.52	9	3.37	45.15	0.08	9	3.37	45.15	0.08	8334.49
achgal															
achist20	157	23.23	7	21.57	55.55	1.68	35	22.64	41.84	1.20	12	22.39	73.39	0.36	10035.03
achnap10	10	7.27	1	7.06	0.40	0.03	1	7.06	0.40	0.03	1	7.06	0.40	0.03	12288.38
achtar10	23	9.63	6	9.11	77.10	1.31	13	50.65	57.40	0.82	22	62.36	93.07	0.76	9095.11
acimdc	2	2.15	1	2.09	0.30	0.01	1	2.09	0.30	0.01	1	2.09	0.30	0.01	7.04
acindn	2	16.43	1	14.94	1.00	0.05	1	14.94	1.00	0.05	1	14.94	1.00	0.05	17.98
acinst	777	14.7	677	8.49	45.90	0.54	376	11.34	27.80	0.57	677	8.49	45.90	0.54	47932.72
ackynu	372	6.06	33	3.58	79.17	2.66	64	4.93	42.74	1.35	70	4.65	78.58	0.34	23353.55
acmbpn	60	8.94	17	7.49	96.56	0.18	8	8.67	49.15	0.48	17	7.49	96.56	0.18	19416.53
acmebz	19	5.15	13	4.10	6.00	0.17	13	4.10	6.00	0.17	13	4.10	6.00	0.17	8056.44
acmtde	794	7.98	525	4.03	110.33	1.61	363	4.91	51.56	1.64	327	5.42	52.00	0.98	38705.03
acnort	184	32.42	7	24.95	68.82	0.84	48	26.41	9.35	0.21	48	26.41	9.35	0.21	33416.84
acnpac10	5	17.48	1	17.35	1.73	0.03	1	17.35	1.73	0.03	1	17.35	1.73	0.03	205.98
acnpec	706	6.00	360	0.91	89.10	2.36	705	2.48	58.15	1.84	145	2.35	73.44	1.21	40618.15
acontn10	53	63.57	4	56.29	61.62	0.45	32	60.47	18.62	0.30	9	60.89	46.26	0.27	10299.99
acpenc10	34	42.83	9	38.79	124.85	1.29	19	39.47	30.71	0.57	19	39.47	30.71	0.57	9806.16
acppca	8	7.20	4	6.51	82.40	0.89	1	6.70	8.87	0.13	1	6.70	8.87	0.13	712.28
acpret03	75	22.26	24	19.76	61.57	1.11	27	21.12	16.02	0.52	12	22.30	55.15	0.29	42830.40
acpyns	124	17.31	59	14.59	66.12	0.86	63	15.46	4.78	0.47	3	15.63	5.48	0.28	16637.39
acrams	13	21.43	7	16.63	118.78	2.02	9	16.83	42.45	1.53	1	16.77	64.85	1.07	4471.35
acsala01	20	4.57	15	4.27	88.76	1.01	15	4.27	88.76	1.01	1	4.49	90.00	0.29	6391.12
acseso10	8	50.75	3	48.84	76.73	0.71	1	49.47	4.27	0.15	1	49.47	4.27	0.15	8387.41
actand	8	21.27	1	19.50	18.77	0.35	1	19.50	18.77	0.35	1	19.50	18.77	0.35	5369.06
acthbz	253	15.88	14	13.65	39.71	0.96	6	13.72	10.45	0.30	6	13.72	10.45	0.30	21824.27
actold	5	6.40	4	5.30	11.69	0.17	4	5.30	11.69	0.17	4	5.30	11.69	0.17	278.00
actysn	76	5.48	27	2.52	67.69	1.33	12	3.31	43.98	1.58	68	4.33	67.69	0.32	12393.50
acurid	74	15.24	22	13.52	126.81	2.23	17	14.51	55.08	0.72	1	14.81	74.80	0.23	16752.70
acvcho	14	2.98	1	2.23	2.62	0.06	1	2.23	2.62	0.06	1	2.23	2.62	0.06	7385.83
acxmol	185	11.41	26	9.15	68.69	0.86	35	9.87	8.53	0.24	35	9.87	8.53	0.24	21763.40
acxmpr	39	10.62	21	8.62	98.98	1.46	14	9.18	19.33	0.31	14	9.18	19.33	0.31	6266.93
acygly11	21	3.55	10	3.07	87.51	1.40	11	3.15	57.03	0.83	1	3.94	90.02	0.05	2006.49
acytid	65	13.19	6	11.93	88.18	0.47	2	12.42	69.32	0.67	1	13.27	83.93	0.27	10510.30
adelox10	325	25.14	6	23.38	49.61	0.38	43	24.06	5.70	0.08	43	24.06	5.70	0.08	34214.26
adenos10	44	48.13	28	46.22	105.72	1.94	25	46.83	8.27	0.60	21	47.73	109.67	0.50	19073.58
adfgip	2	11.44	1	11.34	1.70	0.02	1	11.34	1.70	0.02	1	11.34	1.70	0.02	13.51
adgsmh															
adhela10	47	13.08	18	12.22	82.87	1.03	14	12.70	4.32	0.06	14	12.70	4.32	0.06	6549.88
admann	7	3.09	3	4.09	98.33	0.50	1	4.41	85.94	0.13	1	4.41	85.94	0.13	3204.54
admhep	20	2.44	2	1.93	102.08	0.76	5	3.30	55.75	0.09	5	3.30	55.75	0.09	4486.18
admina	2	30.10	1	25.16	11.90	0.14	1	25.16	11.90	0.14	1	25.16	11.90	0.14	32.25
admopm	443	47.86	95	40.37	112.30	2.00	396	41.94	66.46	1.71	198	42.71	106.37	1.02	40075.98
adrtar	49	6.06	21	3.22	33.89	0.79	21	3.22	33.89	0.79	35	4.78	113.25	0.12	9617.65
adypnl	37	13.31	32	8.22	100.84	2.75	22	14.32	49.91	0.77	22	14.32	49.91	0.77	53052.67
aebdod10	4	16.27	2	15.30	167.40	0.57	1	15.76	0.70	0.07	1	15.76	0.70	0.07	0.33
aenlan															
afcydp	447	14.38	103	9.57	100.53	1.34	44	10.99	74.76	1.38	413	10.91	82.89	0.90	28030.83
afmscy	21	21.74	1	19.91	74.74	0.41	1	19.91	74.74	0.41	2	21.22	95.86	0.12	19407.72
afurpo10	3	17.36	1	16.96	5.40	0.05	1	16.96	5.40	0.05	1	16.96	5.40	0.05	49.44
afutdz10	2	8.77	1	7.59	0.80	0.22	1	7.59	0.80	0.22	1	7.59	0.80	0.22	11.85
afuthu	19	13.04	11	11.27	97.62	0.74	1	11.71	11.77	0.36	1	11.71	11.77	0.36	7967.21
agalam10	103	3.36	6	2.56	66.36	0.49	9	2.90	45.09	0.06	9	2.90	45.09	0.06	9062.01
agluam10	375	4.52	196	2.09	89.45	1.18	276	3.56	9.76	0.22	276	3.56	9.76	0.22	14432.95
aharfu	4	17.95	1	16.68	4.60	0.39	1	16.68	4.60	0.39	1	16.68	4.60	0.39	129.19
ahcdla	2	9.74	1	9.59	0.50	0.01	1	9.59	0.50	0.01	1	9.59	0.50	0.01	23.35
ahditx	6	27.47	2	26.64	84.28	0.13	1	27.40	18.11	0.20	2	26.64	84.28	0.13	1206.05

TABLE II.
Results from the MACROMODEL Monte Carlo Runs.

RESULTS FROM MACROMODEL's MONTE CARLO

NAME	Xraymin		E Best				Tor Best			RMS Best					
	#Confs	Energy	Conf	Energy	TRMS	RMS	Conf	Energy	TRMS	RMS	Conf	Energy	TRMS	RMS	CPUtime
aaxthp															
abaxes	18	218.75	1	217.57	65.04	0.58	2	218.71	0.00	0.00	2	218.71	0.00	0.00	52130.60
abbumo	3	167.06	1	167.02	0.00	0.00	1	167.02	0.00	0.00	1	167.02	0.00	0.00	22708.17
abinor															
abinos	2	-64.13	1	-64.13	0.00	0.00	1	-64.13	0.00	0.00	1	-64.13	0.00	0.00	7708.00
abtoet	343	41.41	2	40.03	96.35	1.36	24	41.38	0.00	0.00	24	41.38	0.00	0.00	85647.11
abztcx	2	163.97	2	163.95	0.00	0.00	2	163.95	0.00	0.00	2	163.95	0.00	0.00	10722.74
acados	417	-85.66	1	-100.57	78.87	0.33	45	-90.56	42.05	0.99	1	-100.57	78.87	0.33	50351.73
acaflr	4	52.15	1	52.14	0.00	0.00	1	52.14	0.00	0.00	1	52.14	0.00	0.00	12685.14
acanil	5	12.84	1	12.84	22.16	0.30	1	12.84	22.16	0.30	1	12.84	22.16	0.30	5092.51
acarap	403	12.15	1	7.62	37.06	0.77	8	12.13	0.00	0.00	8	12.13	0.00	0.00	29806.49
acbnza	24	-2.39	1	-31.81	137.26	1.19	21	-2.40	0.00	0.00	21	-2.40	0.00	0.00	7612.19
acbuol	903	77.82	1	59.05	106.01	2.12	39	68.35	78.61	1.90	13	62.95	79.51	1.26	82186.19
accitr	22	140.65	1	112.59	120.34	0.25	7	140.63	0.00	0.00	7	140.63	0.00	0.00	22938.28
acdxur	213	-82.05	1	-94.18	11.62	0.90	22	-82.06	0.00	0.00	22	-82.06	0.00	0.00	40992.01
acfpch	16	52.92	1	52.71	5.53	0.05	1	52.71	5.53	0.05	1	52.71	5.53	0.05	20628.67
acfucn	103	-27.12	1	-27.54	29.82	0.67	4	-27.13	0.00	0.00	4	-27.13	0.00	0.00	17958.02
acglsp	558	-13.74	1	-18.24	43.22	0.85	11	-13.77	0.00	0.00	11	-13.77	0.00	0.00	57286.40
acglua	107	-116.02	1	-139.67	95.94	1.00	41	-116.02	0.00	0.00	41	-116.02	0.00	0.00	18070.14
achgal	132	-33.16	1	-47.74	61.79	0.67	7	-33.17	0.00	0.00	7	-33.17	0.00	0.00	21861.49
achist	195	-21.70	1	-31.05	102.01	1.44	27	-21.70	0.00	0.00	27	-21.70	0.00	0.00	14162.10
achnap	5	44.17	1	44.16	0.00	0.00	1	44.16	0.00	0.00	1	44.16	0.00	0.00	14451.76
achtar	52	47.19	3	36.78	90.94	0.97	27	47.19	0.05	0.00	27	47.19	0.05	0.00	12772.09
acimdc	4	-82.42	1	-83.93	180.00	0.04	3	-82.42	0.00	0.00	3	-82.42	0.00	0.00	736.37
acindn	9	48.25	1	28.66	155.57	0.90	6	48.24	0.00	0.00	6	48.24	0.00	0.00	4820.71
acinst	265	-13.05	1	-19.64	55.02	0.91	10	-10.56	18.60	0.39	4	-15.68	47.42	0.29	35818.81
ackynu	412	-41.62	1	-47.54	60.27	0.86	31	-40.62	44.31	1.62	21	-41.63	62.61	0.01	29948.58
acmbpn	273	23.44	1	-3.58	74.07	0.90	44	22.98	1.48	0.03	44	22.98	1.48	0.03	33950.20
acmebz	15	-32.36	1	-32.36	0.00	0.00	1	-32.36	0.00	0.00	1	-32.36	0.00	0.00	8755.50
acmtde	712	33.71	1	22.93	79.23	1.08	21	27.95	42.11	1.12	34	28.98	62.29	0.75	26480.00
acnort	75	150.00	1	147.46	67.75	0.84	4	149.96	0.00	0.00	4	149.96	0.00	0.00	63467.33
acnpac	5	15.22	1	15.21	0.00	0.00	1	15.21	0.00	0.00	1	15.21	0.00	0.00	6409.07
acnpec															
acontn	228	405.27	1	396.32	59.52	0.39	22	405.19	0.04	0.00	22	405.19	0.04	0.00	63790.48
acpenc	182	181.30	1	173.32	63.05	1.16	44	179.76	10.03	0.87	13	176.73	64.32	0.65	26835.23
acppca	13	2.41	1	-0.09	117.90	1.19	3	2.41	0.00	0.00	3	2.41	0.00	0.00	9496.05
acpret	246	147.41	1	126.12	63.05	1.20	26	147.36	1.26	0.05	26	147.36	1.26	0.05	99355.68
acpyns	61	-33.30	1	-35.56	60.93	0.70	2	-33.32	0.00	0.00	2	-33.32	0.00	0.00	29822.55
acrams	88	82.52	1	82.49	125.83	0.94	3	82.49	0.27	0.02	3	82.49	0.27	0.02	55859.21
acsala	14	-32.08	1	-34.62	89.21	0.85	3	-32.08	0.00	0.00	3	-32.08	0.00	0.00	8527.09
acseso	9	234.23	1	233.63	66.16	0.62	2	234.21	0.07	0.00	2	234.21	0.07	0.00	38834.27
actand	21	168.52	1	168.48	0.30	0.01	1	168.48	0.30	0.01	1	168.48	0.30	0.01	73507.68
acthbz	201	48.28	1	48.15	79.95	1.08	5	48.26	0.14	0.01	5	48.26	0.14	0.01	33610.48
actold	4	14.98	1	14.97	104.70	0.98	2	14.97	24.70	0.29	2	14.97	24.70	0.29	9891.28
actysn	148	-23.14	1	-27.51	53.11	1.46	28	-23.15	0.00	0.00	28	-23.15	0.00	0.00	21572.26
acurid	255	-78.35	1	-99.63	96.90	2.14	32	-85.06	41.78	0.85	38	-83.84	64.49	0.36	40849.19
acvcho	16	43.70	1	43.69	90.00	0.00	5	51.60	57.81	0.40	1	43.69	90.00	0.00	18989.29
acxmol	48	-1.53	1	-1.55	0.08	0.00	1	-1.55	0.08	0.00	1	-1.55	0.08	0.00	33746.84
acxmpr	156	-62.79	1	-73.22	81.91	1.45	17	-62.79	0.04	0.00	17	-62.79	0.04	0.00	9464.00
acygly	23	-48.59	1	-48.59	0.00	0.00	1	-48.59	0.00	0.00	1	-48.59	0.00	0.00	3661.64
acytid	94	-144.29	1	-161.16	110.19	1.17	29	-144.29	0.15	0.02	29	-144.29	0.15	0.02	30913.19
adelox	68	258.15	1	257.62	78.67	0.35	2	258.11	0.06	0.00	2	258.11	0.06	0.00	51993.27
adenos	208	-85.65	1	-92.57	100.34	0.81	29	-85.65	0.14	0.01	29	-85.65	0.14	0.01	43979.20
adfglp	2	-156.37	1	-156.38	52.40	0.17	1	-156.38	52.40	0.17	1	-156.38	52.40	0.17	4134.40
adgsmh	660	-44.91	1	-44.95	0.03	0.00	1	-44.95	0.03	0.00	1	-44.95	0.03	0.00	77746.09
adhela	76	-23.27	1	-23.27	0.00	0.00	1	-23.27	0.00	0.00	1	-23.27	0.00	0.00	18086.45
admann	31	-80.08	1	-96.11	50.71	0.77	24	-62.13	49.11	0.28	24	-62.13	49.11	0.28	10288.33
admhep	123	-109.36	1	-129.85	61.13	0.79	34	-109.37	0.04	0.00	34	-109.37	0.04	0.00	13637.48
admina	1	-28.52	1	-28.53	0.00	0.00	1	-28.53	0.00	0.00	1	-28.53	0.00	0.00	5818.06
admopm															
adrtar	117	26.46	1	16.83	91.13	1.76	13	25.29	51.65	1.18	8	23.89	119.97	0.39	14846.41
adypnl	557	145.43	1	137.29	114.28	1.59	2	137.29	47.97	1.13	9	137.46	52.42	1.12	101414.07
aebdod	5	111.65	1	109.57	157.00	0.56	2	111.62	0.00	0.00	2	111.62	0.00	0.00	15184.56
aenlan	709	291.27	1	291.23	0.04	0.00	1	291.23	0.04	0.00	1	291.23	0.04	0.00	129229.63
afcydp															
afmscy															
afurpo	2	28.65	1	28.65	0.00	0.00	1	28.65	0.00	0.00	1	28.65	0.00	0.00	4545.20
afutdz															
afuthu	63	-16.08	1	-33.23	69.17	1.30	40	-16.09	0.08	0.01	40	-16.09	0.08	0.01	33102.12
agalam	125	-124.18	1	-139.59	82.22	0.51	21	-124.18	0.04	0.00	21	-124.18	0.04	0.00	18810.92
agluam	342	-71.21	1	-81.20	78.03	1.29	25	-71.21	0.04	0.00	25	-71.21	0.04	0.00	11158.43
aharfu	19	23.45	1	13.35	95.98	0.85	13	23.45	0.00	0.00	13	23.45	0.00	0.00	18218.75
ahcdla	2	97.88	1	97.87	0.00	0.00	1	97.87	0.00	0.00	1	97.87	0.00	0.00	6946.30
ahditx	7	156.57	1	156.39	75.66	0.25	1	156.39	75.66	0.25	1	156.39	75.66	0.25	61380.63

the null value of the heavy atom RMS and the torsional RMS as well as the energy. In fact, using this criterion it is clear from the tables that MACROMODEL-MC succeeded many more times (33) in visiting the reference conformer exactly than SYBYL-RANDOM (0) did. This may simply be the outcome of the force field and minimizer used in these two software suites. Nonetheless, it remains a difference that one must be aware of. It is fortunate that such an example as ACHTAR presented itself in this study because it serves to underscore the stochastic nature of these methods.

Another finding from the data is that there is only a slight correlation between the torsional RMS and the heavy atom RMS values in the SYBYL-RANDOM approach: that is, when one is low it does not follow that the other is also low. Non-quantitative cursory examination of both tables seems to indicate that this trend is true for both methods examined. This finding may be connected to the considerations of symmetry-related atoms in a molecule. This can be seen with all three sets of these numbers that are available from Tables I and II, namely the columns for the best energy conformer, the best torsional RMS conformer, and the best heavy atom RMS conformer. In all cases when the torsional RMS numbers were plotted against the RMS numbers a very noncolinear scatterplot was observed. We have stated in previous articles that both measures taken together are better than either one alone in determining the closeness of a conformer to the reference one. This point is again emphasized with this observation.

The time taken for the SYBYL-RANDOM runs was on average about 14,000 CPU seconds on a SGI ONYX VTX. However, the data were skewed heavily toward the low end of the observed range of 63,884.81 to 0.33 seconds. The number of unique conformers found varied from 0 to 793. Again, the data were heavily skewed toward the low end with no less than 47 molecules having less than 10 conformers found. The MACROMODEL-MC runs were performed on a smaller VAX machine as indicated earlier. Thus the timing of the two runs cannot be directly compared. However, the average time for the runs was roughly 31,000 CPU seconds. This number compares very favorably indeed with the previous one considering the difference between the two machines on which the experiments were run. The data were again skewed toward the low end of the spread from 101414.07 to 736.37, but not as heavily as in the previous case. The number of unique conformers found varied from 0 to 903 with 17 molecules having less than 10 conformers found. MACROMODEL-MC found no acceptable conformers for six molecules while SYBYL-RANDOM found none for four. Again, these were not the same molecules.

No correlations were observed between either the time of the run and the number of rotatable bonds in the molecule or with the number of conformers found and the number of rotatable bonds in the molecule. This is a significant attribute of the so-called stochastic search methods and may need to come into play when considering searches on large, flexible molecules.

Conclusions

It is clear that the savings in time for a stochastic conformational search over a systematic search can be great. However, two considerations need to enter into this claim: (1) the time taken to minimize and compare several conformers to one another is a significant fraction of the total time for the run in both types of searching techniques; and (2) the time used to really complete a run by identifying all of the found low energy conformers the requisite number of times increases as the number of rotatable bonds increases. This is only artificially masked by setting a limit on the number of trials performed. The stochastic method is therefore to be preferred, where only a few low energy conformers are expected. In such cases, the search may be much faster than the systematic search methods. The stochastic searching methods examined did indeed find conformers having energies close to that of the reference conformer and which may thus legitimately be called "low energy conformers."

References

1. A. R. Leach, A survey of methods for searching the conformational space of small and medium-sized molecules, In *Reviews in Computational Chemistry*, vol. II, K. B. Lipkowitz and D. B. Boyd, Eds., VCH, New York, 1991, pp. 1–55.
2. K. Gundertofte, J. Palm, I. Pettersson, and A. Stamvik, *J. Comput. Chem.*, **12**, 200 (1991).
3. H. Boehm, G. Klebe, T. Lorenz, T. Meitzner, L. Siggel, *J. Comput. Chem.*, **11**, 1021 (1990).
4. M. Saunders and W. C. Still, et al. *J. Am. Chem. Soc.*, **112**, 1419 (1990).
5. A. K. Ghose, E. P. Jaeger, P. J. Kowalczyk, M. L. Peterson, and A. M. Treasurywala, *J. Comput. Chem.*, **14**, 1050 (1993).

6. R. S. Judson, E. P. Jaeger, M. L. Peterson, and A. M. Treasurywala, *J. Comput. Chem.*, **14**, 1407 (1993).
7. M. Saunders, *J. Am. Chem. Soc.*, **109**, 3150 (1987).
8. Tripos Associates, Inc., St. Louis, MO.
9. MACROMODEL c/o Prof. W. C. Still, Columbia University, New York, NY.
10. G. Chang, W. C. Guida, and W. C. Still, *J. Am. Chem. Soc.*, **111**, 4379 (1989).
11. M. Clark, R. D. Crammer, and N. van Opdenbosch, *J. Comput. Chem.*, **10**, 982 (1989).
12. S. L. Mayo, B. D. Olafson, and W. A. Goddard, *J. Phys. Chem.*, **94**, 8897 (1990).
13. S. West and K. Davies, *Chem. Des. Auto. News*, **6**, 4 (1991).
14. Unfortunately, a validation of the MACROMODEL force field used here with this standard set of molecules could not be found.